JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 51, 557–576 (1975)

Admissible Relaxation in Variational and Control Problems*

FRANK H. CLARKE

University of British Columbia, Vancouver, B.C. V6T 1W5

Submitted by George Leitmann

1. INTRODUCTION

Consider the following *problem of Bolza* in the calculus of variations: minimize

$$\ell[x(0), x(1)] + \int_0^1 L[t, x(t), \dot{x}(t)]\, dt$$

over a given class of functions $x: [0, 1] \to R^n$. It is frequently preferable to consider instead of this problem the one in which $L(t, s, v)$ is convexified in v. This convexity in v, for example, is necessary for lower-semicontinuity of the integral functional [8] and is important in guaranteeing the existence of a minimum [12] and in deriving necessary conditions [17].

Similar considerations apply to the following *trajectory problem*. Given a mapping E from $[0, 1] \times R^n$ to the subsets of R^n, we call x a trajectory for E if it satisfies the differential inclusion

$$\dot{x}(t) \in E[t, x(t)] \quad \text{a.e.} \quad \text{on} \quad [0, 1].$$

We seek to minimize $\ell[x(0), x(1)]$ over a given class of trajectories for E. In this case, the corresponding convexified problem is the one in which E is replaced by its convex hull.

In either of the above cases, the passage to the new convexified problem is called *relaxation*, and the new problem is called *relaxed*. Within the context of the problem of Bolza, relaxation is akin to admitting the "generalized curves" of Young [18]. Relaxation of the less classical trajectory problem has also been treated, as we shall see.

A natural question to ask is the following: Under what conditions is the

* The contents of this article are taken from the author's doctoral dissertation, University of Washington, June 1973, written under the supervision of R. T. Rockafellar.

relaxed problem equivalent to the original one? It is known that in general a solution to one need not be a solution to the other. We prove an "admissible relaxation" theorem for slightly generalized forms of each of the above problems. These theorems give conditions under which a solution to the original problem is necessarily a solution to the relaxed problem, both problems having the same minimum. These theorems are a form of necessary condition, for they assert that in order that a function solve the problem it is necessary that the function also be a relaxed solution. In fact, we show that Theorem 1, which deals with relaxing the problem of Bolza, has as an immediate consequence a generalization of the necessary condition of Weierstrass [13] in the classical calculus of variations.

A special case of the trajectory problem, in which $E(t, s)$ has the form $f(t, s, U)$ for some function f and set U, has been treated by Warga and by Varaiya. The relaxation theorems they obtain [15, Theorem 2.2; 16, Theorem 2.3] are subsumed by our Theorem 2.

Sections 3 and 4 deal with the problem of Bolza and the trajectory problem, respectively, while Section 2 obtains some preliminary results on set-valued mappings and their trajectories. Two of these results (Propositions 2 and 3) are very basic in nature. They are the basis of many existence proofs for trajectories, and forms of them have been obtained under various sets of hypotheses by other authors (see [6, Theorems 2.1, 2.2; 5, Theorems 1, 3]). We believe the versions we give to be the most general.

An interesting feature of the proofs is the use that is made of the fact that a problem of Bolza may be recast as a trajectory problem, and vice versa. The first relaxation result obtained is for a certain special trajectory problem (Proposition 4); this then figures in the proof of Theorem 1 concerning the problem of Bolza. In turn, Theorem 1 is used to prove Theorem 2 on relaxing trajectory problems.

2. Differential Inclusions and Trajectories

We are given a closed *multifunction* E: $[0, 1] \times R^n \to R^n$; i.e. a function mapping $[0, 1] \times R^n$ to the closed subsets of R^n. An *arc* is an absolutely continuous function x: $[0, 1] \to R^n$, and a trajectory for E is an arc satisfying

$$\dot{x}(t) \in E[t, x(t)] \quad \text{a.e.} \tag{1}$$

("a.e." will denote "for almost all t in [0, 1]").

A particularly important example of a multifunction arises in optimal control. Let f: $[0, 1] \times R^n \times U \to R^n$ be given, and define

$$E(t, s) = \{f(t, s, u) \colon u \in U\}. \tag{2}$$

Then under mild hypotheses, x is a trajectory for this E iff there exists a measurable function $u(t)$ taking values in U such that

$$\dot{x}(t) = f[t, x(t), u(t)] \quad \text{a.e.}$$

(This is often known as Filippov's lemma.) For this reason, we shall sometimes refer to an extremum problem with a constraint like (1) as a control problem.

The multifunction E is *measurable* in t if, for each x in R^n and closed subset C of R^n, the set

$$\{t \in [0, 1]: E(t, x) \cap C \neq \phi\}$$

is Lebesgue measurable.

DEFINITION 1. Let W be a subset of $[0, 1] \times R^n$. We say E is *Lipschitz* in x on W if there exists an integrable function $k(t)$ on $[0, 1]$ such that, given any (t, x_1) and (t, x_2) in W, and v_1 in $E(t, x_1)$, there is some v_2 in $E(t, x_2)$ satisfying

$$| v_1 - v_2 | \leqslant k(t) \, | x_1 - x_2 | .$$

Since E is closed-valued this is equivalent to a Lipschitz condition relative to the Hausdorff metric. When E has the form in (2), E has the above property if f is Lipschitz in x.

We denote Euclidean distance by d.

DEFINITION 2. The function ρ: $[0, 1] \times R^n \times R^n \to [0, \infty]$ is defined by

$$\rho(t, x, v) = d[v, E(t, x)].$$

Whenever $E(t, s)$ is nonempty, ρ is uniformly Lipschitz in v with Lipschitz constant 1. As we now see, it is the presence or absence of Lipschitz behavior in x that determines whether or not E is Lipschitz in x.

PROPOSITION 1. *The following are equivalent:*

(a) *E is Lipschitz in x on W, with Lipschitz function $k(t)$.*

(b) *For any (t, x_1) and (t, x_2) in W, and for any v_1 and v_2 in R^n,*

$$| \rho(t, x_1, v_1) - \rho(t, x_2, v_2) | \leqslant k(t) \, | x_1 - x_2 | + | v_1 - v_2 | .$$

(c) *For any (t, x_1) and (t, x_2) in W and v in $E(t, x_1)$ or $E(t, x_2)$,*

$$| \rho(t, x_1, v) - \rho(t, x_2, v) | \leqslant k(t) \, | x_1 - x_2 | .$$

Proof. Suppose (a), and let t, x_1, x_2, v_1, and v_2 as in (b) be given. Choose ζ_1 in $E(t, x_1)$ so that

$$|\zeta_1 - v_1| = \rho(t, x_1, v_1).$$

By hypothesis there is some ζ_2 in $E(t, x_2)$ such that

$$|\zeta_1 - \zeta_2| \leqslant k(t) |x_1 - x_2|.$$

Then we have

$$\begin{aligned}\rho(t, x_2, v_2) &\leqslant \rho(t, x_2, \zeta_1) + |\zeta_1 - v_2| \\ &\leqslant |\zeta_1 - \zeta_2| + |\zeta_1 - v_1| + |v_1 - v_2| \\ &\leqslant k(t) |x_1 - x_2| + \rho(t, x_1, v_1) + |v_1 - v_2|.\end{aligned}$$

Since we may obtain a similar inequality with (x_1, v_1) and (x_2, v_2) reversed, (b) follows.

Evidently (b) implies (c), so we need only show that (a) follows from (c). Let (t, x_1) and (t, x_2) in W be given, as well as v_1 in $E(t, x_1)$. Then from (c),

$$\rho(t, x_2, v_1) \leqslant \rho(t, x_1, v_1) + k(t) |x_1 - x_2| = k(t) |x_1 - x_2|,$$

whence some v_2 in $E(t, x_2)$ exists which satisfies

$$|v_1 - v_2| \leqslant k(t) |x_1 - x_2|,$$

and (a) follows. Q.E.D.

It is not difficult to show that $\rho[t, x(t), \dot{x}(t)]$ is a measurable function of t when E is measurable in t and Lipschitz in x on W, and x is an arc such that $[t, x(t)]$ lies in W (see [2, Proposition 3.9]).

Definition 3. For E and x as above, define

$$d(x) = \int_0^1 \rho[t, x(t), \dot{x}(t)]\, dt.$$

We may regard $d(x)$ as a measure of how close x is to being a trajectory for E; indeed, x is a trajectory iff $d(x) = 0$. The following result states roughly that an arc x is as close in the sup norm to a trajectory as it is to being a trajectory. ($B(\epsilon, s)$ is an open ball of radius ϵ about s; $\| x \|$ is the sup norm of the arc x.)

Proposition 2. *Let E be measurable in t and Lipschitz in x on W. There exists a positive number K with the following property:*

Given any arc x and ϵ in $(0, \infty]$ such that $\{t\} \times B[\epsilon, x(t)] \subset W$ for all t in $[0, 1]$ and $d(x) < \epsilon/K$, then there exists a trajectory y for E with $x(0) = y(0)$ such that

$$\| x - y \| \leqslant Kd(x).$$

We omit the proof [2, p. 45], since it is essentially the same as that of Theorem 1 in [5]. The main difference is the behavior of E in t; in [5] continuity is assumed. This is used only in a measurable selection lemma [5, Lemma 4], and recent selection results [1] show that measurability in t will suffice.

DEFINITION 4. The arc x is a *relaxed trajectory* for E iff

$$\dot{x}(t) \in \operatorname{co} E[t, x(t)] \quad \text{a.e.},$$

where "co" denotes convex hull. For contrast, we shall sometimes refer to a trajectory as an *original* trajectory. (Note: $\operatorname{co} E$ is not necessarily closed-valued although E is.)

DEFINITION 5. E is *integrably bounded* on W if there is an integrable function $\gamma(t)$ on $[0, 1]$ such that for all (t, x) in W, for all v in $E(t, x)$, $| v | \leqslant \gamma(t)$.

DEFINITION 6. The arc x is said to lie in W if, for every t, $[t, x(t)]$ lies in W. We say that W is *open in the* sup *norm* if, for every arc x in W, there is a positive ϵ such that (t, s) lies in W for every s within ϵ of $x(t)$.

PROPOSITION 3. *Let E be measurable in t, Lipschitz in x on W, and integrably bounded on W, where W is open in the* sup *norm. Then for every relaxed trajectory y in W, given any positive δ there is a trajectory x with $x(0) = y(0)$ and $\| x - y \| < \delta$.*

Proof. We may assume that $\{t\} \times B[\lambda, y(t)]$ is contained in W for all t. Let k be the Lipschitz function for E as in Definition 1, and let K be the constant whose existence is asserted in Proposition 2, for $\epsilon = \lambda/2$. Choose a positive α such that

$$\alpha < \min \left\{ \epsilon \middle/ \left[K \int_0^1 k(t)\, dt \right], \delta \middle/ \left[1 + K \int_0^1 k(t)\, dt \right], \epsilon \right\}.$$

Let γ be the function of Definition 5, and choose a positive integer m

such that for any subinterval I of $[0, 1]$ with measure no greater than $1/m$,

$$\int_I \gamma(t)\,dt < \alpha/2.$$

Let I_j be the interval $[(j-1)/m, j/m)$, $i = 1, 2,\ldots, m$. Because $E[t, y(t)]$ is measurable and integrably bounded, we may apply a theorem of Aumann [7, p. 143] to deduce the existence of integrable functions f_j such that $f_j(t) \in E[t, y(t)]$ for almost all t in I_j, and

$$\int_{I_j} f_j(t)\,dt = \int_{I_j} \dot{y}(t)\,dt \qquad (j = 1, 2,\ldots, m).$$

Let f be the function which is equal to f_j on I_j, and define the arc x_0 by

$$x_0(t) = y(0) + \int_0^t f(\tau)\,d\tau.$$

We then have

$$| x_0(t) - y(t)| = \int_0^t [f(\tau) - \dot{y}(\tau)]\,d\tau \leqslant \int_{I_j} | f(\tau) - \dot{y}(\tau)|\,d\tau \qquad \text{(for some } j)$$

$$\leqslant \int_{I_j} 2\gamma(\tau)\,d\tau < \alpha.$$

It follows that $\| x_0 - y \| \leqslant \alpha < \epsilon$ and hence that $\{t\} \times B[\epsilon, x_0(t)]$ is contained in W for all t. We also derive, since $\dot{x}_0(t) \in E[t, y(t)]$ a.e.,

$$\rho[t, x_0(t), \dot{x}_0(t)] \leqslant k(t)\,| x_0(t) - y(t)| \leqslant k(t)\,\alpha \quad \text{a.e.,}$$

whence

$$d(x_0) \leqslant \alpha \int_0^1 k(t)\,dt < \epsilon/K.$$

From Proposition 2 we conclude that a trajectory x exists such that $x(0) = x_0(0)$ and $\| x - x_0 \| \leqslant Kd(x_0)$. Then

$$\| x - y \| \leqslant \| x - x_0 \| + \| x_0 - y \| \leqslant Kd(x_0) + \alpha \leqslant K\alpha \int_0^1 k(t)\,dt + \alpha$$

$$= \alpha \left[1 + K \int_0^1 k(t)\,dt\right] < \delta.$$

Q.E.D.

Consider now the following variational problem: to minimize $f(x)$ over the arcs x which are trajectories for E, for which $x(0)$ lies in C_0, and which are in W (here f is a real-valued function defined on the class of arcs and C_0

is a given subset of R^n). We will use the following type of notation to symbolize such a problem:

$$\text{minimize}\{f(x): x \text{ in } W,\ x(0) \in C_0,\ \dot{x} \in E(t, x) \text{ a.e.}\}. \tag{3}$$

If we replace E in the above problem by co E, the resulting problem is called the relaxed problem, in keeping with Definition 4. In either case, if the minimum exists and is attained at a feasible arc x (i.e., an arc satisfying the constraints), then x is said to solve the problem. Evidently the minimum in the relaxed problem is no greater than that in the original problem.

PROPOSITION 4. *Let the arc z solve problem* (3), *where f is continuous in the* sup *norm, E is measurable in t, Lipschitz in x on W, and integrably bounded on W. If W is open in the* sup *norm, z also solves the relaxed problem:*

$$\text{minimize}\{f(x): x \text{ in } W,\ x(0) \in C_0,\ \dot{x} \in \text{co } E(t, x) \text{ a.e.}\}.$$

Proof. Suppose the result false. Then there is a relaxed trajectory y in W with $y(0) \in C_0$ and $f(y) < f(z)$. But y may be approximated in the sup norm to any degree of closeness by trajectories x with $x(0) = y(0)$ by Proposition 3. Hence, by the continuity of f there is a trajectory x interior to W with $x(0) \in C_0$ such that $f(x) < f(z)$. This is the required contradiction. Q.E.D.

3. THE PROBLEM OF BOLZA

We are given functions $\ell: R^n \times R^n \to (-\infty, \infty]$, $L: [0, 1] \times R^n \times R^n \to R$ and a real-valued function f desired on the class of arcs. We consider the following problem of Bolza in the calculus of variations: to minimize

$$f(x) + \ell[x(0), x(1)] + \int_0^1 L[t, x(t), \dot{x}(t)]\, dt \tag{4}$$

over a subclass of all arcs x. The fact that ℓ is allowed to attain the value $+\infty$ greatly increases the versatility of this problem. For example, we may take ℓ to be the indicator of a set C; i.e. the function which equals 0 on C and $+\infty$ elsewhere. Then minimizing (4) is equivalent to minimizing

$$f(x) + \int_0^1 L[t, x(t), \dot{x}(t)]\, dt$$

over a subclass of arcs x which satisfy the constraint

$$[x(0), x(1)] \in C.$$

The versatility of the problem of Bolza is also discussed in [3, 4, 11, 12].

DEFINITION 7. Let z be an arc, and let ϵ be a positive number. The tube of radius ϵ about z, denoted $S(\epsilon, z)$, is the class of all arcs x such that $|x(t) - z(t)| < \epsilon$ for all t in $[0, 1]$.

As before, W will denote a subset of $[0, 1] \times R^n$. We continue to use the language of Definition 6. Note that W is open in the sup norm if, for any arc x in W, there is some positive ϵ such that every arc of $S(\epsilon, x)$ also lies in W.

We denote by $\tilde{L}(t, s, v)$ the convexification of L in v. That is, for each (t, s), $\tilde{L}(t, s, \cdot)$ is the largest convex function majorized by $L(t, s, \cdot)$ (in [10] this is called the convex hull of $L(t, s, \cdot)$). In general, $\tilde{L}$ may have to be $-\infty$.

We denote by A^∞ the class of arcs whose derivatives are essentially bounded.

THEOREM 1. *Let z solve the problem*

$$\text{minimize } \left\{ f(x) + \ell[x(0), x(1)] + \int_0^1 L(t, x, \dot{x})\, dt \colon x \text{ in } W,\ x \in z + A^\infty \right\},$$

where f is Lipschitz in the sup *norm on W, $\ell: R^n \times R^n \to (-\infty, \infty]$ is l.s.c. (lower semicontinuous), L is measurable in t, and W is open in the* sup *norm. Suppose also that for each positive r there exists an integrable function k_r such that when (t, s_1), (t, s_2) lie in W, and whenever v_1 and v_2 lie within r of $\dot{z}(t)$,*

$$|L(t, s_1, v_1) - L(t, s_2, v_2)| \leqslant k_r(t)\, |(s_1 - s_2, v_1 - v_2)| .$$

Then $\tilde{L}[t, z(t), \dot{z}(t)]$ is finite a.e., and

$$\tilde{L}[t, z(t), \dot{z}(t)] = L[t, z(t), \dot{z}(t)] \quad \text{a.e.} \tag{5}$$

Moreover, if x is in W and in $z + A^\infty$, and if $\ell[x(0), x(1)] < \infty$, then $\int_0^1 \tilde{L}(t, x, \dot{x})\, dt$ is well-defined and finite, and

$$\begin{aligned} f(x) + \ell[x(0), x(1)] &+ \int_0^1 \tilde{L}(t, x, \dot{x})\, dt \\ &\geqslant f(z) + \ell[z(0), z(1)] + \int_0^1 L(t, z, \dot{z})\, dt. \end{aligned} \tag{6}$$

REMARKS. The proof is postponed until some consequences of the theorem are discussed. Inequality (6) above implies that z solves a relaxed (i.e., convexified) problem. Simpler statements are possible when ℓ has a special form, as we show below. Equality (5) is the essence of the necessary condition of Weierstrass, as we now show.

COROLLARY 1. (Generalized Weierstrass condition). *Under the hypotheses of Theorem* 1, *there exists for almost all t an element $p(t)$ of R^n such that for all $v \in R^n$,*

$$L[t, z(t), \dot{z}(t) + v] - L[t, z(t), \dot{z}(t)] \geqslant v \cdot p(t). \tag{7}$$

If $L[t, z(t), \cdot]$ admits a gradient at $\dot{z}(t)$, then $p(t) = \nabla_v L[t, z(t), \dot{z}(t)]$.

Proof. For almost all t, $\tilde{L}[t, z(t), \cdot]$ is a convex function finite at $\dot{z}(t)$. Furthermore, $\tilde{L}[t, z(t), \cdot]$ does not assume the value $+\infty$, since L is finite-valued. It follows from convex analysis that $\tilde{L}[t, z(t), \cdot]$ has a subgradient at $\dot{z}(t)$; i.e. a vector $p(t)$ such that (7) holds for the function $\tilde{L}$ instead of L. But since, for all v,

$$\tilde{L}[t, z(t), \dot{z}(t) + v] \leqslant L[t, z(t), \dot{z}(t) + v]$$

and equality holds at $v = 0$, we easily derive (7) itself. The last assertion follows readily. Q.E.D.

Remark. In the classical calculus of variations, the Lipschitz hypothesis of Theorem 1 follows from the hypothesis that L is C^1, and the condition (7) (with $\nabla_v L$ instead of $p(t)$) is often stated in terms of the Weierstrass "excess function." In that setting z is usually assumed to solve the problem of Bolza with respect to the arcs with continuous derivatives, for example, rather than the arcs in $z + A^\infty$ but the above results hold unchanged in such cases [2].

COROLLARY 2. *Let z solve the problem*

$$\text{minimize } \left\{ f(x) + \int_0^1 L(t, x, \dot{x})\, dt : x \in S(\epsilon, z),\ x \in z + A^\infty,\ [x(0), x(1)] \in C \right\},$$

where f and L satisfy the hypotheses of Theorem 1. *Then z also solves the relaxed problem*

$$\text{minimize } \left\{ f(x) + \int_0^1 \tilde{L}(t, x, \dot{x})\, dt : x \in S(\epsilon, z),\ x \in z + A^\infty,\ [x(0), x(1)] \in C \right\}.$$

COROLLARY 3. *Let z, f, L, and ℓ be as in Theorem* 1, *where in addition ℓ is finite-valued. Then z solves the relaxed problem*

$$\text{minimize}\left\{f(x)+\ell[x(0),x(1)]+\int_0^1 \tilde{L}(t,x,\dot{x})\,dt\colon x \text{ in } W,\ x\in z+A^\infty\right\}.$$

We now proceed to the proof, beginning with the following preliminary result.

LEMMA 1. *Let the arc z solve the problem*

$$\text{minimize}\left\{f(x)+\ell[x(0),x(1)]+\int_0^1 L(t,x,\dot{x})\,dt\colon x \text{ in } W,\ |\dot{x}-\dot{z}|\leqslant r\right\},$$

where f is continuous in the sup *norm, ℓ is continuous, W is open in the* sup *norm, and*

$$\sup\{|s-z(t)|\colon (t,s)\in W\}=\epsilon<\infty.$$

Suppose further that L is measurable in t and that there exists an integrable function $k(t)$ such that, for each t, $L(t,\cdot,\cdot)$ is Lipschitz on the set

$$\{(s,v)\colon (t,s)\in W,\ |v-\dot{z}(t)|\leqslant r\},$$

with Lipschitz constant $k(t)$. Let $L_r(t,s,\cdot)$ denote the function $L(t,s,\cdot)$ restricted to those v satisfying $|v-\dot{z}(t)|\leqslant r$, and let $\tilde{L}_r(t,s,\cdot)$ denote the convexification of $L_r(t,s,\cdot)$. (Equivalently, let $L_r(t,s,v)=L(t,s,v)$ for $|v-\dot{z}(t)|\leqslant r$ and $+\infty$ otherwise; then $\tilde{L}_r$ is the convexification of L_r in v.) Then z solves the following problem:

$$\text{minimize}\left\{f(x)+\ell[x(0),x(1)]+\int_0^1 \tilde{L}_r(t,x,\dot{x})\,dt\colon x \text{ in } W,\ |\dot{x}-\dot{z}|\leqslant r\right\},$$

and

$$\tilde{L}_r(t,z,\dot{z})=L(t,z,\dot{z})\quad \text{a.e.}$$

Proof. We shall adopt the following convention: s^*, v^*, etc., will denote points of R^{n+1}, where we shall consider having adjoined an extra coordinate s^0 to a point s in R^n, etc. Thus $s^*=(s^0,s)$. Similarly, an arc x^* to R^{n+1} is of the form (x^0,x), where x^0 is a real-valued arc and x is, as before, R^n-valued.

Using this convention, define the multifunction E by:

$$E(t,s^*)=\{v^*\colon L(t,s,v)\leqslant v^0\leqslant L(t,z,\dot{z})+k(t)(r+\epsilon),\ |v-\dot{z}|\leqslant r\}.$$

Let W^* be the set

$$\{(t,s^*)\colon (t,s)\in W\},$$

and define the arc z^* by

$$z^*(t) = \left\{ \int_0^t L[\tau, z(\tau), \dot{z}(\tau)]\, d\tau, z(t) \right\}.$$

One may easily verify that z^* solves the problem:

minimize$\{f(x) + \ell[x(0), x(1)] + x^0(1) \colon x^* \text{ in } W^*, x^0(0) = 0, \dot{x}^* \in E(t, x^*)\text{a.e.}\}$.

It follows also that E is measurable in t, closed-valued, and integrably bounded on W^*.

Claim. E is Lipschitz in s^* on W^*.

Note first that the Lipschitz hypothesis on L implies that E is nonempty on W^*. Let $(t, s_1{}^*)$ and $(t, s_2{}^*)$ in W^* and $v_1{}^*$ in $E(t, s_1{}^*)$ be given. We must produce some $v_2{}^*$ in $E(t, s_2{}^*)$ such that

$$| v_1{}^* - v_2{}^* | \leqslant k(t) \, | s_1{}^* - s_2{}^* | .$$

If $L(t, s_2, v_1) \leqslant v_1{}^0$, we may take $v_2{}^* = v_1{}^*$. If $L(t, s_2, v_1) > v_1{}^0$, put $v_2 = v_1$ and $v_2{}^0 = L(t, s_2, v_1)$. Note then that $v_2{}^*$ lies in $E(t, s_2{}^*)$ and

$$\begin{aligned} | v_1{}^* - v_2{}^* | &= | v_1{}^0 - L(t, s_2, v_1)| \\ &= L(t, s_2, v_1) - v_1{}^0 \leqslant L(t, s_2, v_1) - L(t, s_1, v_1) \\ &\leqslant k(t) \, | s_1 - s_2 | \leqslant k(t) \, | s_1{}^* - s_2{}^* | , \end{aligned}$$

which proves the claim.

We now have the requisites to apply Proposition 4, and we conclude that z^* solves the relaxed problem:

minimize$\{f(x) + \ell[x(0), x(1)] + x^0(1) \colon x^* \text{ in } W^*, \dot{x}^* \in \operatorname{co} E(t, x^*) \text{ a.e.}\}$.

Because of the Lipschitz hypothesis, whenever (t, s^*) lies in W^*, $E(t, s^*)$ contains all points $[L(t, s, v), v]$ with $| v - \dot{z}(t)| \leqslant r$, and hence, for (t, s^*) in W^*,

$$\operatorname{co} E(t, s^*) = \{v^* \colon \tilde{L}_r(t, s, v) \leqslant v^0 \leqslant L(t, z, \dot{z}) + k(t)\,(r + \epsilon)\}.$$

A translation of terms yields the fact that z solves the stated relaxed problem, with the minimum being the same as in the original problem (the measurability of $\tilde{L}_r$ follows from [10, Corollary 17.15]).

We have then

$$\int_0^1 \tilde{L}_r(t, z, \dot{z})\, dt = \int_0^1 L(t, z, \dot{z})\, dt,$$

and combining this with the fact that $\tilde{L}_r \leqslant L$ (when $\tilde{L}_r$ is finite), we deduce the final equality of the lemma. Q.E.D.

Proof of Theorem 1. Using the same asterisk convention as in the proof of Lemma 1, we define the arc z^* by

$$z^*(t) = \{\ell[z(0), z(1)], z(t)\}.$$

Let C be the closed set

$$\{(s_0{}^*, s_1{}^*): \ell(s_0, s_1) \leqslant s_1{}^0\},$$

and define $L^*: [0, 1] \times R^{n+1} \times R^{n+1} \to R$ by

$$L^*(t, s^*, v^*) = L(t, s, v).$$

For any positive ϵ, define

$$W_\epsilon = \{(t, s) \in W: \{t\} \times B(\epsilon, s) \subset W\}.$$

(Note that any arc x in W will be contained in W_ϵ for ϵ sufficiently small, since W is open in the sup norm.) Define also, for each t in $[0, 1]$,

$$M_\epsilon(t) = \{(s^*, v^*): (t, s) \in W_\epsilon, \mid s^* - z^*(t)| < 1/\epsilon, \mid v^* - \dot{z}^*(t)| \leqslant 1/\epsilon\}.$$

It follows that for all ϵ sufficiently small (so that z lies in W_ϵ), z^* solves the problem:

$$\text{minimize} \Big\{ f(x) + x^0(1) + \int_0^1 L^*(t, x^*, \dot{x}^*)\, dt: (x^*, \dot{x}^*) \in M_{\epsilon/2}(t), [x^*(0), x^*(1)] \in C \Big\}.$$

A simple proof by contradiction (see [2, Lemma 2.8; or 4, Lemma 3.2]) yields the fact that, for some positive m, z^* solves the problem:

$$\text{minimize} \Big\{ f(x) + x^0(1) + m d_C[x^*(0), x^*(1)] + \int_0^1 L^*(t, x^*, \dot{x}^*)\, dt: (x^*, \dot{x}^*) \in M_\epsilon(t) \Big\}.$$

We conclude from this that z^* solves the problem:

$$\text{minimize}\left\{f(x) + x^0(1) + md_C[x^*(0), x^*(1)] \right.$$
$$\left. + \int_0^1 L^*(t, x^*, \dot{x}^*)\, dt\colon x^* \text{ in } W_\epsilon^*, \mid \dot{x}^* - \dot{z}^* \mid \leqslant 1/\epsilon\right\},$$

where

$$W_\epsilon^* = \{(t, s^*)\colon (t, s) \in W_\epsilon, \mid s^0 - z^0(t)| < 1/\epsilon\}.$$

The conditions of Lemma 1 (with $r = 1/\epsilon$) are met, and we deduce that z^* solves the problem:

$$\text{minimize}\left\{f(x) + x^0(1) + md_C[x^*(0), x^*(1)] \right.$$
$$\left. + \int_0^1 \tilde{L}_r^*(t, x^*, \dot{x}^*)\, dt\colon x^* \text{ in } W_\epsilon^*, \mid \dot{x}^* - \dot{z}^* \mid \leqslant r\right\},$$

and that

$$\tilde{L}_r^*(t, z^*, \dot{z}^*) = L^*(t, z^*, \dot{z}^*) = L(t, z, \dot{z}) \quad \text{a.e.}$$

It is easy to see that $\tilde{L}_r^*(t, z^*, \dot{z}^*) = \tilde{L}_r(t, z, \dot{z})$, whence

$$\tilde{L}_r(t, z, \dot{z}) = L(t, z, \dot{z}) \quad \text{a.e.} \tag{8}$$

Returning to a problem on R^n, we may conclude from the above that z solves the problem:

$$\text{minimize}\left\{f(x) + \ell[x(0), x(1)] \right.$$
$$\left. + \int_0^1 \tilde{L}_r(t, x, \dot{x})\, dt\colon x \text{ in } W_\epsilon, \mid \ell[x(0), x(1)] - \ell[z(0), z(1)]| < r\right\}.$$

It follows from [10, Corollary 17.1.5] that $\tilde{L}_r \downarrow \tilde{L}$ as $r \uparrow \infty$ (or as $\epsilon \downarrow 0$), which, along with (8), yields (5) of the theorem.

Now let x be any arc as described in the theorem. For $\epsilon = 1/r$ sufficiently small we have x in W_ϵ, $\mid \dot{x} - \dot{z} \mid \leqslant r$ a.e., and

$$\mid \ell[x(0), x(1)] - \ell[z(0), z(1)]| < r.$$

From the fact that z solves the above problem we conclude

$$f(x) + \ell[x(0), x(1)] + \int_0^1 \tilde{L}_r(t, x, \dot{x})\, dt \geqslant f(z) + \ell[z(0), z(1)] + \int_0^1 L(t, z, \dot{z})\, dt$$

(it follows from the Lipschitz hypothesis that the integral on the left is finite). Recalling that $\tilde{L}_r \downarrow \tilde{L}$, and invoking Lebesgue's monotone convergence theorem, we derive (6). Q.E.D.

4. Relaxing the Trajectory Problem

We shall apply Theorem 1 to obtain a relaxation theorem for the following *trajectory problem:*

$$\text{minimize}\{f(x) + \ell[x(0), x(1)]: \dot{x}(t) \in E[t, x(t)] \text{ a.e., } x \text{ in } W\}.$$

Definition 8. Define functions $\phi^i: R^n \to [-\infty, \infty]$ $(i = 0, 1)$ as follows:

$$\phi^0(s) = \inf\{f(x) + \ell[x(0) + s, x(1)]: \dot{x} \in E(t, x) \text{ a.e., } x \text{ in } W\},$$
$$\phi^1(s) = \inf\{f(x) + \ell[x(0), x(1) + s]: \dot{x} \in E(t, x) \text{ a.e., } x \text{ in } W\}.$$

The trajectory problem is *calm* if, for $i = 0$ *or* 1, we have

$$\liminf_{s \to 0} [\phi^i(s) - \phi^i(0)]/|s| > -\infty.$$

Theorem 2. *Let z solve the above trajectory problem, where*

$$\ell: R^n \times R^n \to (-\infty, \infty]$$

is l.s.c., W is open in the sup *norm, f is Lipschitz in the* sup *norm on W, and E is Lipschitz in s on W. Then, if the problem is calm, z also solves the relaxed problem:*

$$\text{minimize}\{f(x) + \ell[x(0), x(1)]: \dot{x}(t) \in \overline{\text{co}}\, E[t, x(t)] \text{ a.e., } x \text{ in } W\}.$$

Remark. We may handle the case when z is a local solution by putting $W = S(\epsilon, z)$. The above result states that any nonrelaxable problem must be ill-posed (i.e., unstable, not calm) in the following sense: Arbitrarily small perturbations of the problem produce proportionally unbounded variations in the minimum. Such problems could be expected to pose serious computational difficulties; related discussions may be found in [9, 14].

Theorem 2 raises the question of when problems are calm. An obvious sufficient condition for calmness is for ℓ to be Lipschitz near $[z(0), z(1)]$, in either variable. Other conditions were established in [2], and a more complete discussion will appear elsewhere. The importance of calmness is increased by another of its implications (see [2, 3]); the necessary conditions corresponding to calm problems may be taken *normal*, i.e., a certain multiplier is nonzero. This fact is of particular interest in applying the maximum principle of optimal control theory.

An example of a nonrelaxable (and hence noncalm) optimal control problem is given in [16, p. 118]. One may also produce noncalm problems that are relaxable.

Proof of the theorem. We remark first that we may reduce to the case where

$$\liminf_{s\to 0}[\phi^1(s) - \phi^1(0)]/|s| > -\infty,$$

by the following artifice: set $\hat{E}(t, s) = -E(1-t, s)$, $\hat{x}(t) = x(1-t)$, $\hat{\ell}(s_0, s_1) = \ell(s_1, s_0)$, and define

$$\hat{W} = \{(t, s): (1-t, s) \in W\}.$$

Then consider the problem:

$$\text{minimize}\{f(x) + \hat{\ell}[x(0), x(1)]: \dot{x} \in \hat{E}(t, x) \text{ a.e., } x \text{ in } \hat{W}\}.$$

This problem continues to satisfy the hypotheses of Theorem 2, and if the original problem satisfied the calmness condition relative to ϕ^0, the new one is calm relative to ϕ^1. But relaxing this new problem is equivalent to relaxing the original (note for example that $\dot{x} \in E(t, x)$ is equivalent to $\dot{\hat{x}} \in \hat{E}(t, \hat{x})$). We thus assume that the inequality in Definition 8 holds for $i = 1$.

We shall use a convention similar to that used earlier, but now the notation s^* will denote a point (s^1, s^2, s^3) of $R^n \times R^n \times R$. Thus an arc x^* has components x^1 and x^2 mapping $[0, 1]$ to R^n and x^3 mapping $[0, 1]$ to R, etc. We shall reformulate the problem to one on R^{2n+1}. For positive ϵ, define

$$\begin{aligned}
E^*(t, s^*) &= E(t, s^1) \times \{0\} \times [0, \infty),\\
z^*(t) &= \{z(t), z(1), \ell[z(0), z(1)]\},\\
C_0 &= \{s^*: \ell(s^1, s^2) \leqslant s^3\}\\
C_1 &= \{s^*: s^1 = s^2\}\\
W_\epsilon^* &= \{(t, s^*): (t, s^1) \in W_\epsilon, |s^* - z^*(t)| < 1/\epsilon\},
\end{aligned}$$

where W_ϵ is defined as in the proof of Theorem 1.

Note that any trajectory x^* for E^* with $x^*(0) \in C_0$ satisfies

$$x^2(0) = x^2(1), \qquad x^3(1) \geqslant x^3(0). \tag{9}$$

Let ϵ be sufficiently small so that z is in W_ϵ. We claim that z^* solves the following problem (for which it is clearly feasible):

$$\text{minimize}\{f(x^1) + x^3(1): \dot{x}^* \in E^*(t, x^*), x^*(0) \in C_0, x^*(1) \in C_1, x^* \text{ in } W_\epsilon^*\}.$$

If this were false, there would be an admissible arc x^* with

$$f(x^1) + x^3(1) < f(z^1) + z^3(1).$$

But then x^1 is a trajectory for E in W, and

$$\begin{aligned} &f(x^1) + \ell[x^1(0), x^1(1)] \\ &\quad = f(x^1) + \ell[x^1(0), x^2(1)] \\ &\quad = f(x^1) + \ell[x^1(0), x^2(0)] \leqslant f(x^1) + x^3(0) \leqslant f(x^1) + x^3(1) < f(z^1) + z^3(1) \\ &\quad = f(z) + \ell[z(0), z(1)], \end{aligned}$$

which contradicts the optimality of z.

LEMMA 2. *There exists a number r such that z^* solves the problem:*

$$\text{minimize}\{f(x^1) + x^3(1) + rd_{C_1}[x^*(1)] + r^2 d^*(x^*)\colon x^*(0) \in C_0\,,\ x^* \text{ in } W_\epsilon^*\},$$

where d^* corresponds to E^* as in Definition 2.

Proof. Suppose not. Then for each n there is an arc x_n^* in W_ϵ^* with $x_n^*(0) \in C_0$ such that

$$f(x_n^1) + x_n^3(1) + nd_{C_1}[x_n^*(1)] + n^2 d^*(x_n^*) < f(z) + z^3(1) = \phi^1(0). \tag{10}$$

Since the first two terms on the left are bounded, both $d_{C_1}[x_n^*(1)]$ and $d^*(x_n^*)$ converge to 0 as n goes to ∞. Let K be the constant of Proposition 2. For all large n we have

$$d^*(x_n^*) < \epsilon/K,$$

and hence by Proposition 2 there exist trajectories y_n^* for E^* with $y_n^*(0) = x_n^*(0)$ and

$$\| x_n^* - y_n^* \| \leqslant Kd^*(x_n^*). \tag{11}$$

It follows that $d_{C_1}[y_n^*(1)]$ also converges to 0, and hence that

$$s_n = y_n^2(1) - y_n^1(1)$$

converges to 0 as well. We also note

$$| s_n | \leqslant 2d_{C_1}[y_n^*(1)]. \tag{12}$$

Letting M be a Lipschitz constant for f, we have

$$\begin{aligned}
\phi^1(s_n) &\leqslant f(y_n^1) + \ell[y_n^1(0), y_n^1(1) + s_n] \\
&= f(y_n^1) + \ell[y_n^1(0), y_n^2(1)] = f(y_n^1) + \ell[y_n^1(0), y_n^2(0)] \qquad \text{(by (9))} \\
&\leqslant f(y_n^1) + y_n^3(1) \\
&\leqslant f(x_n^1) + x_n^3(1) + M \| x_n^1 - y_n^1 \| + \| x_n^3 - y_n^3 \| \\
&\leqslant \phi^1(0) - n d_{C_1}[x_n^*(1)] - n^2 d^*(x_n^*) + (M+1) \| x_n^* - y_n^* \| \\
&\qquad \text{(by (10))} \\
&\leqslant \phi^1(0) - n d_{C_1}[x_n^*(1)] + [(M+1) K - n^2]\, d^*(x_n^*) \qquad \text{(by (11))} \\
&\leqslant \phi^1(0) - n d_{C_1}[y_n^*(1)] + n \| x_n^* - y_n^* \| + [(M+1) K - n^2]\, d^*(x_n^*) \\
&\leqslant \phi^1(0) - n d_{C_1}[y_n^*(1)] + [[n + M + 1] K - n^2]\, d^*(x_n^*) \\
&\leqslant \phi^1(0) - n d_{C_1}[y_n^*(1)] \qquad \text{(for } n \text{ large)} \\
&\leqslant \phi^1(0) - (n/2) \mid s_n \mid \qquad \text{(by (12)).}
\end{aligned}$$

We derive

$$[\phi^1(s_n) - \phi^1(0)] / \mid s_n \mid \leqslant -n/2,$$

and since $s_n \to 0$ we contradict the calmness assumption. This completes the proof of the lemma.

Recalling the equality

$$d^*(x^*) = \int_0^1 \rho(t, x^*, \dot{x}^*)\, dt,$$

we see that Theorem 1 may be applied to the problem of Lemma 2. The conclusion is that z^* solves the problem:

$$\text{minimize} \left\{ f(x^1) + x^3(1) + r d_{C_1}[x^*(1)] \right.$$
$$\left. + \int_0^1 r^2 \tilde{\rho}^*(t, x^*, \dot{x}^*)\, dt \colon x^*(0) \in C_0,\ x^* \text{ in } W_\epsilon^*,\ x^* \in z^* + A^\infty \right\}. \tag{13}$$

Note that $\tilde{\rho}^*$ is nonnegative, since ρ^* is.

LEMMA 3. *Let $\hat{\rho}$ be the distance function corresponding to $\overline{\text{co}}\, E^*$; i.e.,*

$$\hat{\rho}(t, s^*, v^*) = d[v^*, \overline{\text{co}}\, E^*(t, s^*)].$$

Then $\hat{\rho} = \tilde{\rho}^*$.

Proof. Fix t and s^*. It is easy to verify that $\hat{\rho}(t, s^*, \cdot)$ is convex and no greater than $\rho^*(t, s^*, \cdot)$, whence

$$\hat{\rho}(t, s^*, v^*) \leqslant \tilde{\rho}^*(t, s^*, v^*)$$

for all v^*, by definition of $\tilde{\rho}^*$. Now fix v^*. From [10, Corollary 17.1.5] we have

$$\tilde{\rho}^*(t, s^*, v^*) = \inf(\Sigma\lambda_i\rho^*(t, s^*, v_i^*): \Sigma\lambda_i v_i^* = v^*\},$$

where $\{\lambda_i\}$ is a convex combination. There is a point ζ in $\overline{\text{co}}\, E^*(t, s^*)$ such that

$$\hat{\rho}(t, s^*, v^*) = | v^* - \zeta | .$$

By Caratheodory's Theorem there exist a finite convex combination $\{\lambda_i\}$ and points ζ_i in $E^*(t, s^*)$ such that

$$\zeta = \Sigma\lambda_i\zeta_i .$$

Then

$$\begin{aligned}\tilde{\rho}^*(t, s^*, v^*) &= \tilde{\rho}^*[t, s^*, \Sigma\lambda_i(v^* + \zeta_i - \zeta)]\\ &\leqslant \Sigma\lambda_i\tilde{\rho}^*(t, s^*, v^* + \zeta_i - \zeta)\end{aligned}$$

(since $\tilde{\rho}^*(t, s^*, \cdot)$ is convex)

$$\leqslant \Sigma\lambda_i |(v^* + \zeta_i - \zeta) - \zeta_i |$$

(since $\zeta_i \in E^*(t, s^*)$)

$$= | v^* - \zeta | = \hat{\rho}(t, s^*, v^*).$$

Thus $\tilde{\rho}^* \leqslant \hat{\rho}$, and the lemma is proved. It is easy to see that $\overline{\text{co}}\, E^*$ is Lipschitz in s^* on W_ϵ^*, since E^* is, and hence $\tilde{\rho}^*$ satisfies (b) of Proposition 1 on W_ϵ^*. Because of this, it follows easily that z^* solves the problem:

$$\text{minimize}\Big\{f(x^1) + x^3(1) + rd_{C_1}[x^*(1)] + \int_0^1 r^2\tilde{\rho}^*(t, x^*, \dot{x}^*)\, dt: x^*(0) \in C_0\, , x^* \text{ in } W_\epsilon^*\Big\} .$$

(This is just problem (10) with the constraint $x^* \in z^* + A^\infty$ removed. The essential point in justifying this removal, along with the above Lipschitz property of $\tilde{\rho}^*$, is that for any arc x^* there is an arc y^* in $z^* + A^\infty$ with both $\| x^* - y^* \|$ and $\int | \dot{x}^* - \dot{y}^* | \, dt$ arbitrarily small.)

We may now conclude that z^* solves:

$$\text{minimize}\{f(x^1) + x^3(1) \colon x^*(0) \in C_0\,,\ x^*(1) \in C_1\,,$$
$$\dot{x}^* \in \overline{\text{co}}\, E^*(t, x^*) \text{ a.e., } x^* \text{ in } W_\epsilon{}^*\}.$$

Because of the obvious equality

$$\overline{\text{co}}\, E^*(t, s^*) = \overline{\text{co}}\, E(t, s^1) \times \{0\} \times [0, \infty),$$

and since ϵ is arbitrarily small, we return finally to R^n and conclude that z solves the relaxed problem:

$$\text{minimize}\{f(x) + \ell[x(0), x(1)] \colon \dot{x} \in \overline{\text{co}}\, E(t, x) \text{ a.e., } x \text{ in } W\}.$$

Q.E.D.

References

1. C. Castaing, Sur les multi-applications mesurables, *Rev. Inf. Rech. Op.* **1** (1967) 91–126.
2. F. H. Clarke, Necessary conditions for nonsmooth problems in optimal control and calculus of variations, dissertation, University of Washington, June 1973.
3. F. H. Clarke, Necessary conditions for nonsmooth variational problems, *in* "Proceedings of the Fourteenth Biennial Seminar of the Canadian Mathematical Congress," Lecture Notes in Economics and Mathematical Systems 106, pp. 70–91, Springer–Verlag, New York, 1974.
4. F. H. Clarke, The Euler–Lagrange differential inclusion, *J. Differential Equations*, to appear.
5. A. F. Filippov, Classical solutions of differential equations with multivalued right-hand side, *SIAM J. Control* **5** (1967), 609–621.
6. H. Hermes, Existence and properties of solutions of $\dot{x} \in R(t, x)$, *SIAM Studies Appl. Math.* **5** (1969).
7. H. Hermes, On the structure of attainable sets for generalized differential equations and control systems, *J. Differential Equations* **9** (1971), 141–154.
8. C. Olech, A necessary and sufficient condition for lower semicontinuity of certain integral functionals, to appear.
9. E. Polak, An historical survey of computational methods in optimal control, *SIAM Rev.* **15** (1973), 572–576.
10. R. T. Rockafellar, "Convex Analysis," Princeton Press, Princeton, N. J., 1970.
11. R. T. Rockafellar, Conjugate convex functions in optimal control and the calculus of variations, *J. Math. Anal. Appl.* **32** (1970), 174–222.
12. R. T. Rockafellar, Existence theorems for general control problems of Bolza and Lagrange, *Advances in Math.* **15** (1975), 312–333.
13. H. Sagan, "Introduction to the Calculus of Variations," McGraw–Hill, New York, 1969.

14. A. M. STEINBERG, Application of relaxed solutions to minimum sensitivity optimal control, *J. Optimization Theory Appl.* **10** (1972), 178–186.
15. P. VARAIYA, On the trajectories of a differential system, *in* "Mathematical Theory of Control," Academic Press, New York, 1967.
16. J. WARGA, Relaxed variational problems, *J. Math. Anal. Appl.* **4** (1962), 111–128.
17. J. WARGA, Necessary conditions for minimum in relaxed variational problems, *J. Math. Anal. Appl.* **4** (1962), 129–145.
18. L. C. YOUNG, "Calculus of Variations and Optimal Control Theory," Saunders, Philadelphia, 1969.